Intel® Itanium™ Processor Core

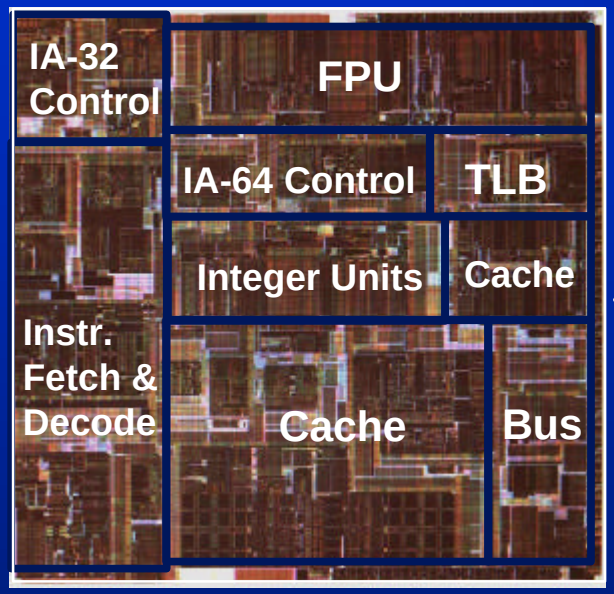
Harsh Sharangpani

Principal Engineer and IA-64 Microarchitecture Manager

Intel Corporation

intel®

Itanium™ Processor Silicon



Core Processor Die

4 x 1MB L3 cache

Machine Characteristics

Frequency	800 MHz
Transistor Count	25.4M CPU; 295M L3
Process	0.18u CMOS, 6 metal layer
Package	Organic Land Grid Array
Machine Width	6 insts/clock (4 ALU/MM, 2 Ld/St, 2 FP, 3 Br)
Registers	14 ported 128 GR & 128 FR; 64 Predicates
Speculation	32 entry ALAT, Exception Deferral
Branch Prediction	Multilevel 4-stage Prediction Hierarchy
FP Compute Bandwidth	3.2 GFlops (DP/EP); 6.4 GFlops (SP)
Memory -> FP Bandwidth	4 DP (8 SP) operands/clock
Virtual Memory Support	64 entry ITLB, 32/96 2-level DTLB, VHPT
L2/L1 Cache	Dual ported 96K Unified & 16KD; 16KI
L2/L1 Latency	6 / 2 clocks
L3 Cache	4MB, 4-way s.a., BW of 12.8 GB/sec;
System Bus	2.1 GB/sec; 4-way Glueless MP Scalable to large (512+ proc) systems

EPIC compared to Dynamic Scheduled RISC

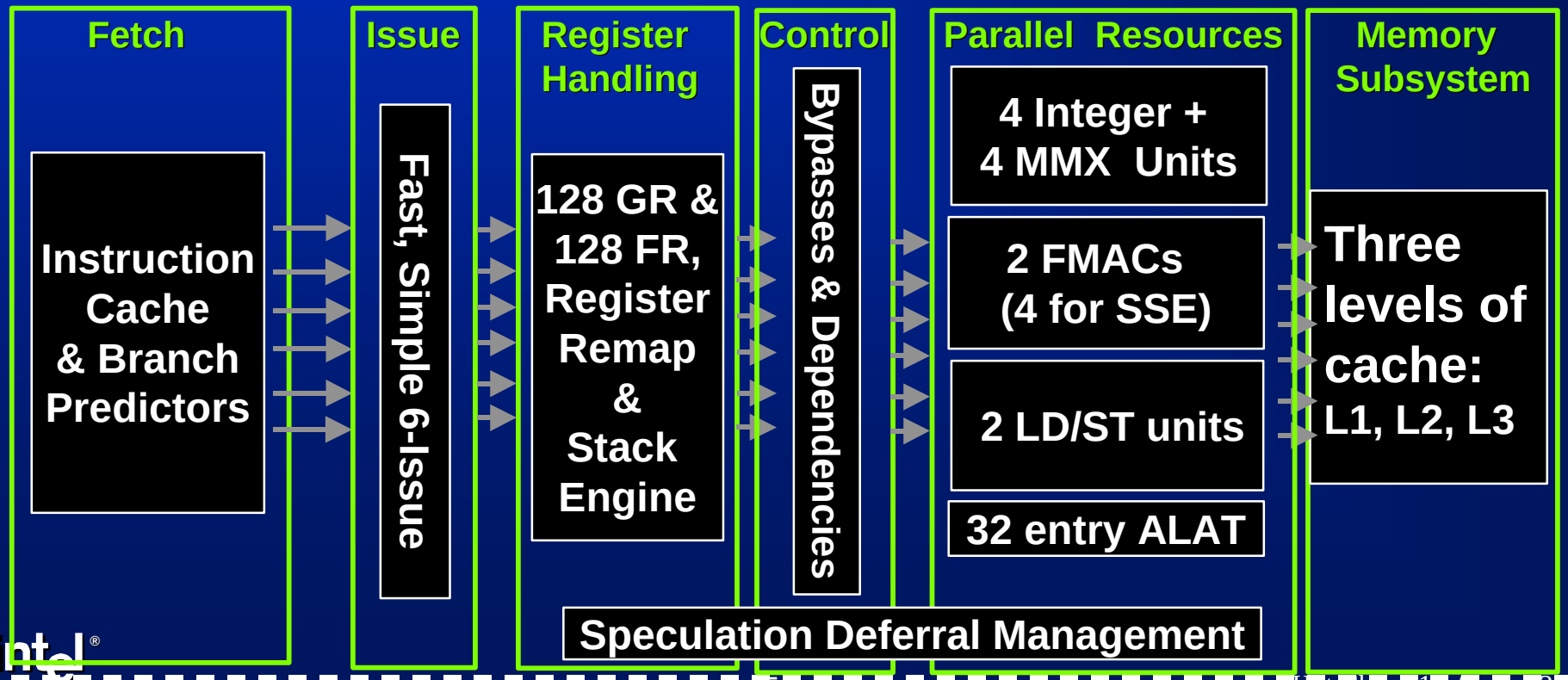
Bottleneck	Itanium EPIC Approach	Dynamic RISC Approach
Scheduling Scope	Entire compilation scope	Traditional compiler + limited hardware window
Memory Latency & Control Flow Barriers	Control Speculation across compiler scope; Data Speculation for undisambiguated memory; Extensive Memory Hints	Hardware Scheduling across dynamic window assisted by Memory Order Buffer
Control Flow Disruptions	Predication for flaky branches; Extensive branch/prefetch Hints; Superscalar branching;	Large Dynamic Branch Predictors; 1 branch/clock.
Operand Delivery	Large Register File, with Stacking & Rotation	Small Architectural File with Register Rename Tables
Interprocedural Overhead	Stacking for parameter passing	Require spill/fill to memory or registers

Itanium™ EPIC Design Maximizes SW-HW Synergy

Architecture Features programmed by compiler:

Branch Hints ***Explicit Parallelism*** ***Register Stack & Rotation*** ***Predication*** ***Data & Control Speculation*** ***Memory Hints***

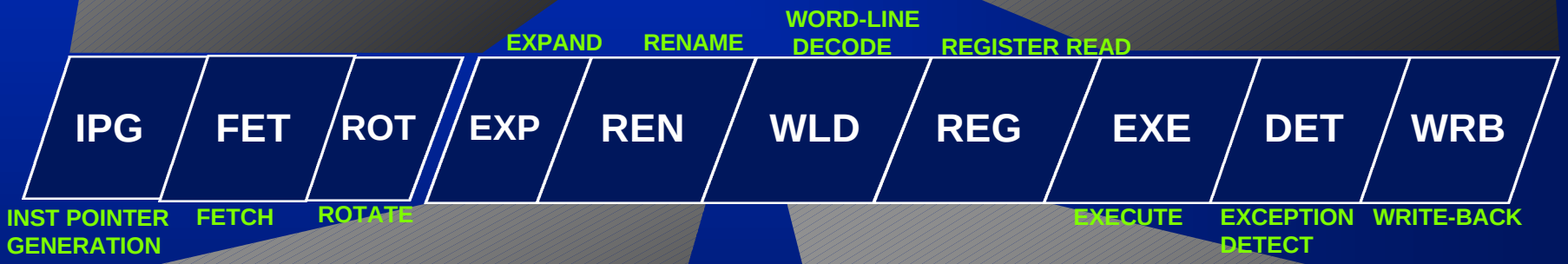
Micro-architecture Features in hardware:



10 Stage In-Order Core Pipeline

Front End
- Pre-fetch/Fetch of up to 6 instructions/cycle
- Hierarchy of branch predictors
- Decoupling buffer

Execution
- 4 single cycle ALUs, 2 ld/str
- Advanced load control
- Predicate delivery & branch
- Nat/Exception//Retirement

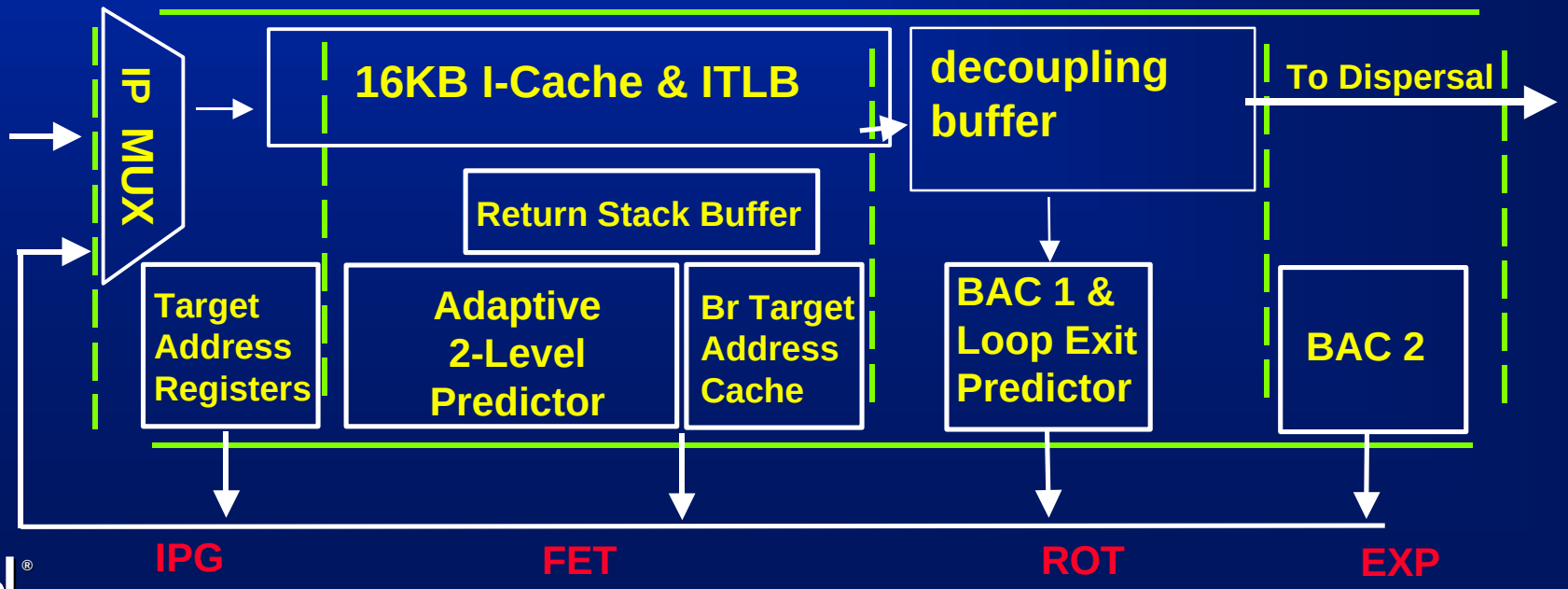


Instruction Delivery
- Dispersal of up to 6 instructions on 9 ports
- Reg. remapping
- Reg. stack engine

Operand Delivery
- Reg read + Bypasses
- Register scoreboard
- Predicated dependencies

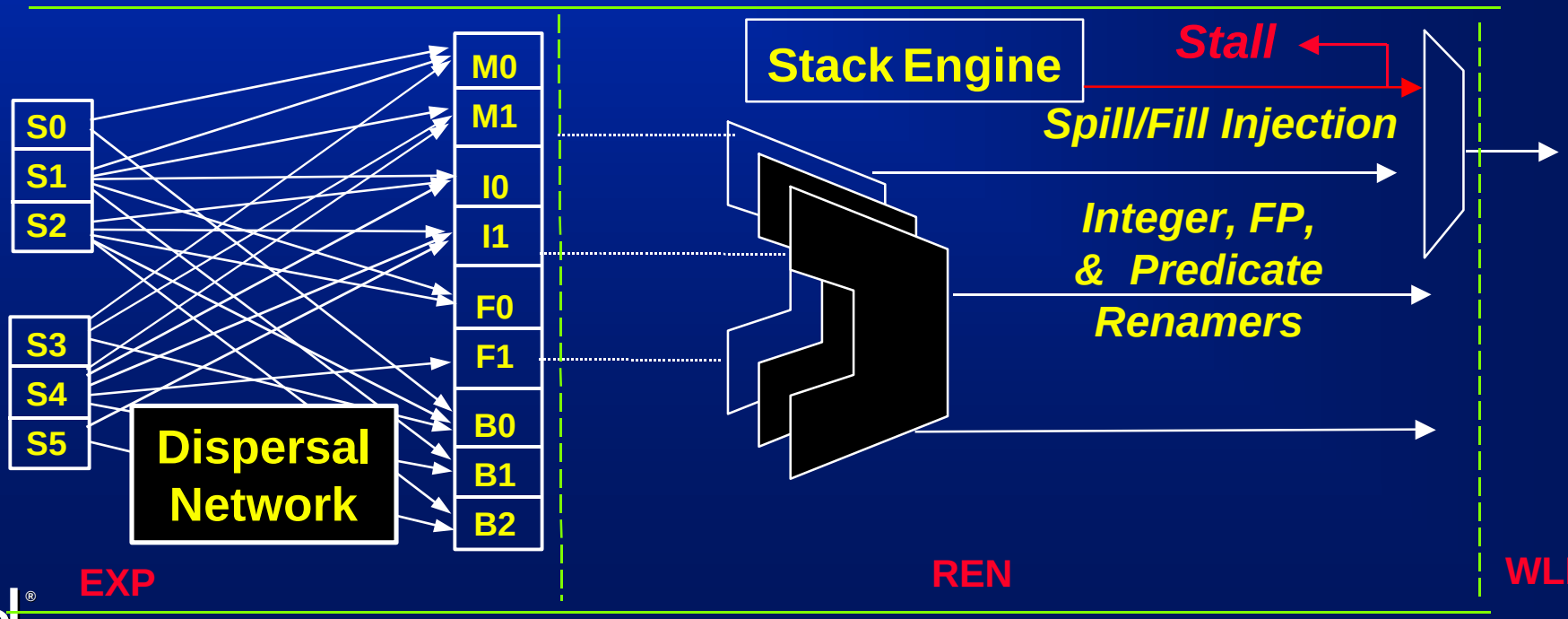
Front End

- SW-triggered prefetch loads target code early using BRP hints
- I-Fetch of 32 Bytes/clock feeds an 8-bundle decoupling buffer
- Branch hints combine with predictor hierarchy to improve branch prediction, delivering upto four progressive resteers
  - 4 TARs under compiler control.
  - Adaptive 2-level predictor (512-entry 2-way + 64-entry Multiway); 64-entry Target Address Cache fed by hints; Return stack buffer;
  - Perfect loop-exit predictor, BAC1, BAC2



Instruction Delivery

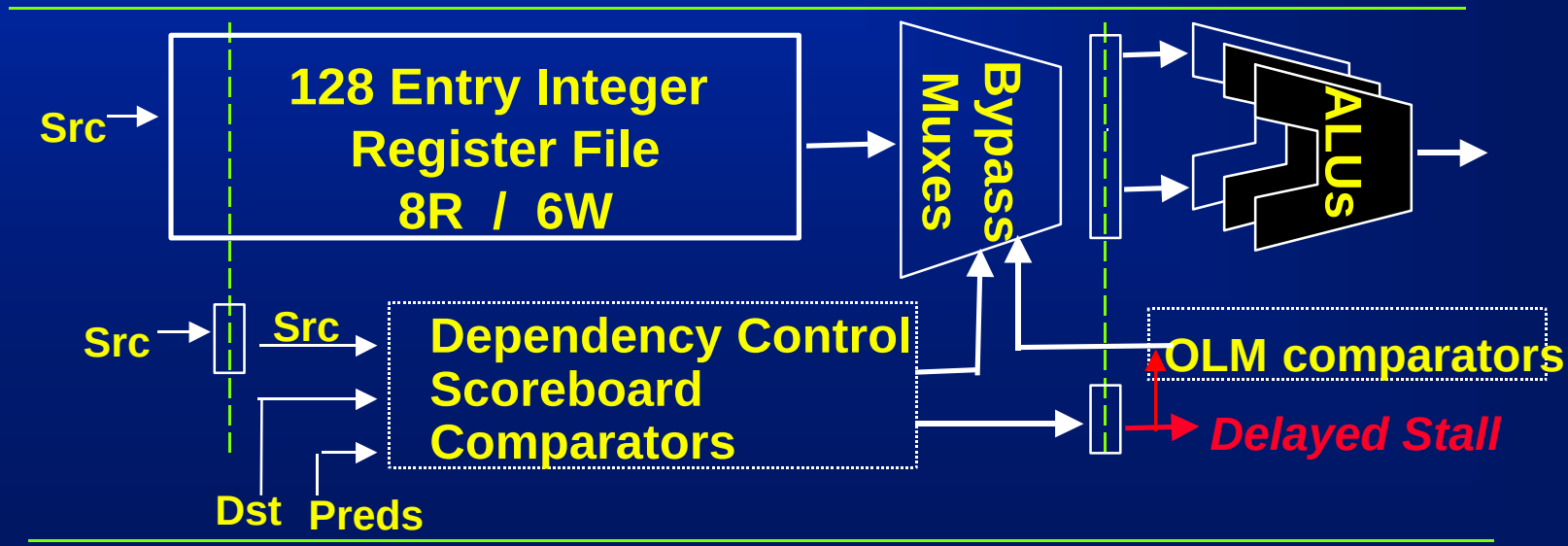
- Stop bits eliminate dependency checking; Templates simplify routing; 1st available dispersal from 6 syllables to 9 issue ports
- Stacking eliminates most register spill / fills
  - Register remapping done via several parallel 7-bit adders
  - Stack engine performs the few required spill/fills
- REN stage supports renaming for stacking & rotation



Operand Delivery

- Multiported register file + mux hierarchy delivers operands in REG
- Unique "*Delayed Stall*" mechanism used for register dependencies
  - Avoids pipeline flush or replay on unavailable data
  - Stall computed in REG, but core pipeline stalls only in EXE
  - Special Operand Latch Manipulation (OLM) captures data returns into operand latches, to mimic register file read

➔ Retains benefits of "stall paradigm" on wide and hi-frequency machine



Execution Resources

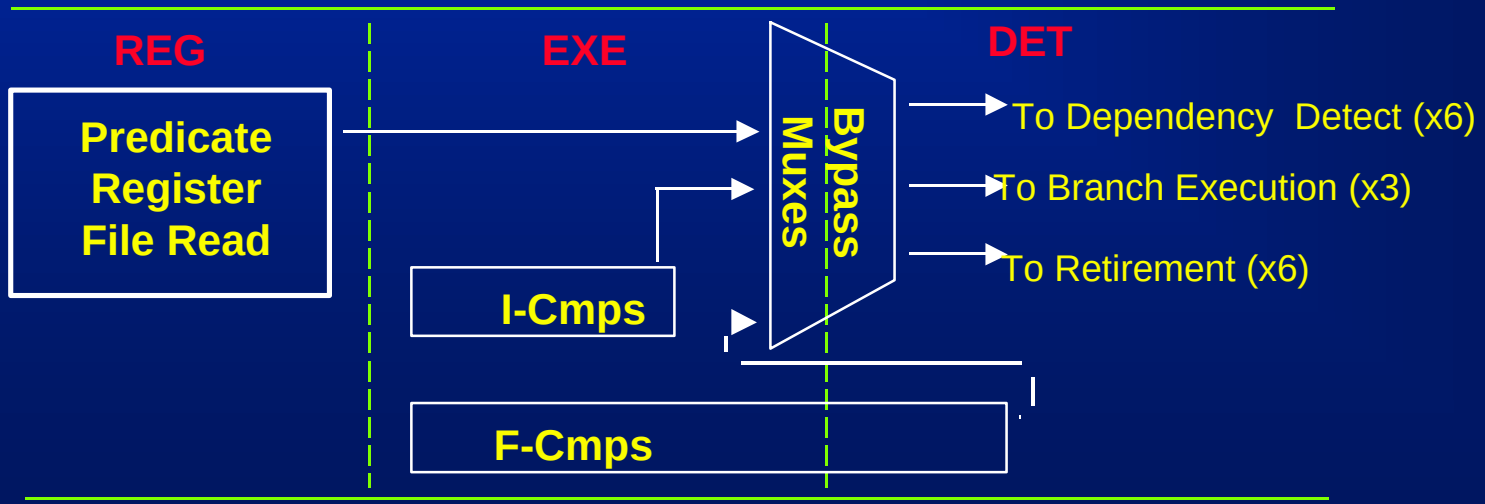
Memory and Integer Resources: Instruction Class	Ports				Latency (clocks)
	M0	M1	I0	I1	
ALU (Add, shift-add, logical, addp4, cmp)	•	•	•	•	1
Sign/zero extend, MoveLong			•	•	1
Fixed Extract/Deposit, TBit, TNaT			•		1
Multimedia ALU	•	•	•	•	2
MM Shift, Avg, Mix, Pack			•	•	2
Move to/from BR/PR/ARs, Packed Multiply, PopCount			•		2
LD/ST/Prefetch/SetF/Cache Control	•	•			2+
Memory Mngmt/System/GetF	•				2+

FP Resources: Instruction Class	Ports		Latency (clocks)
	F0	F1	
FMAC, SIMD FMAC	•	•	5
Fixed Multiply	•	•	7
Fset, Fchk	•	•	1
FCompare	•		2
FP Logicals/Class/Min/Max	•		5

Branch Resources: Instruction Class	Ports		
	B0	B1	B2
Cond/Uncond	•	•	•
Call/Ret/Indirect	•	•	•
Branch.iA, EPC	•	•	•
Loop, BSW, Cover			•
RFI			•

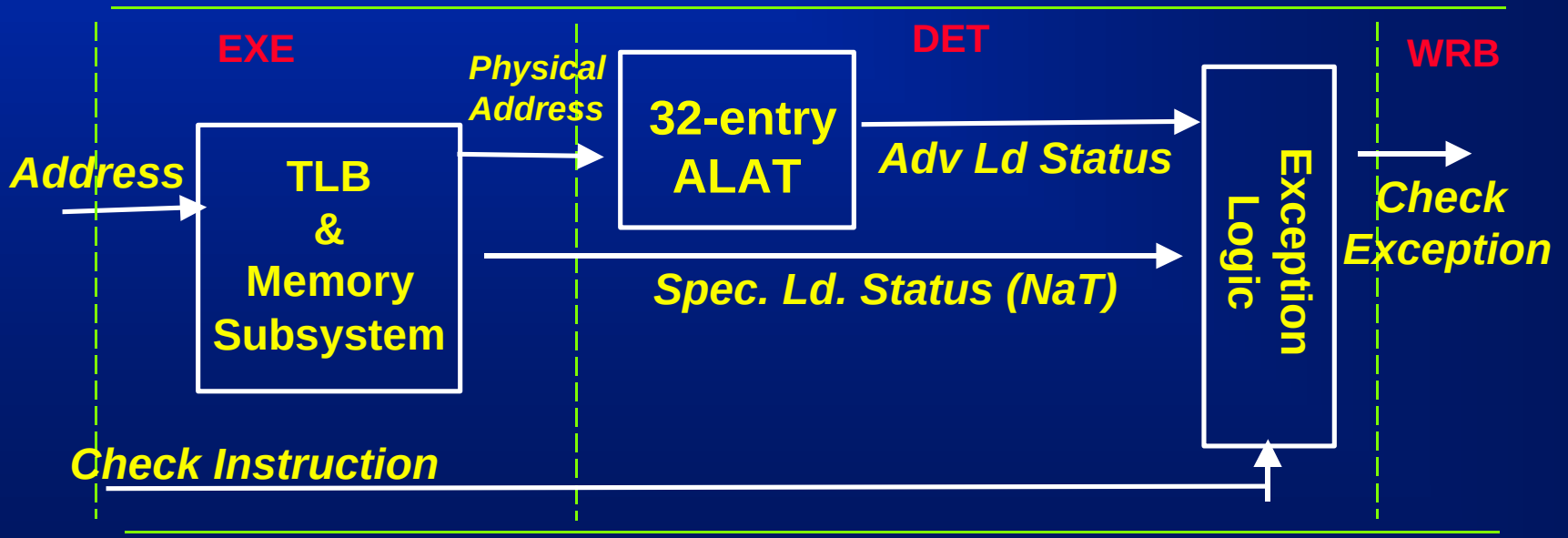
Predication Support

- Basic strategy: All instructions read operands and execute
  - Canceled at retirement if predicates off
- Predicates generated in EXE (by cmps), delivered in DET, & feed into retirement, branch execution and dependency detection
- Smart control cancels false stalls on predicated dependencies
  - Special detection exists in REG for cancelled producer or consumer
- Predication supported transparently - branches (& mispredicts) eliminated without introduction of spurious stalls



Speculation Hardware

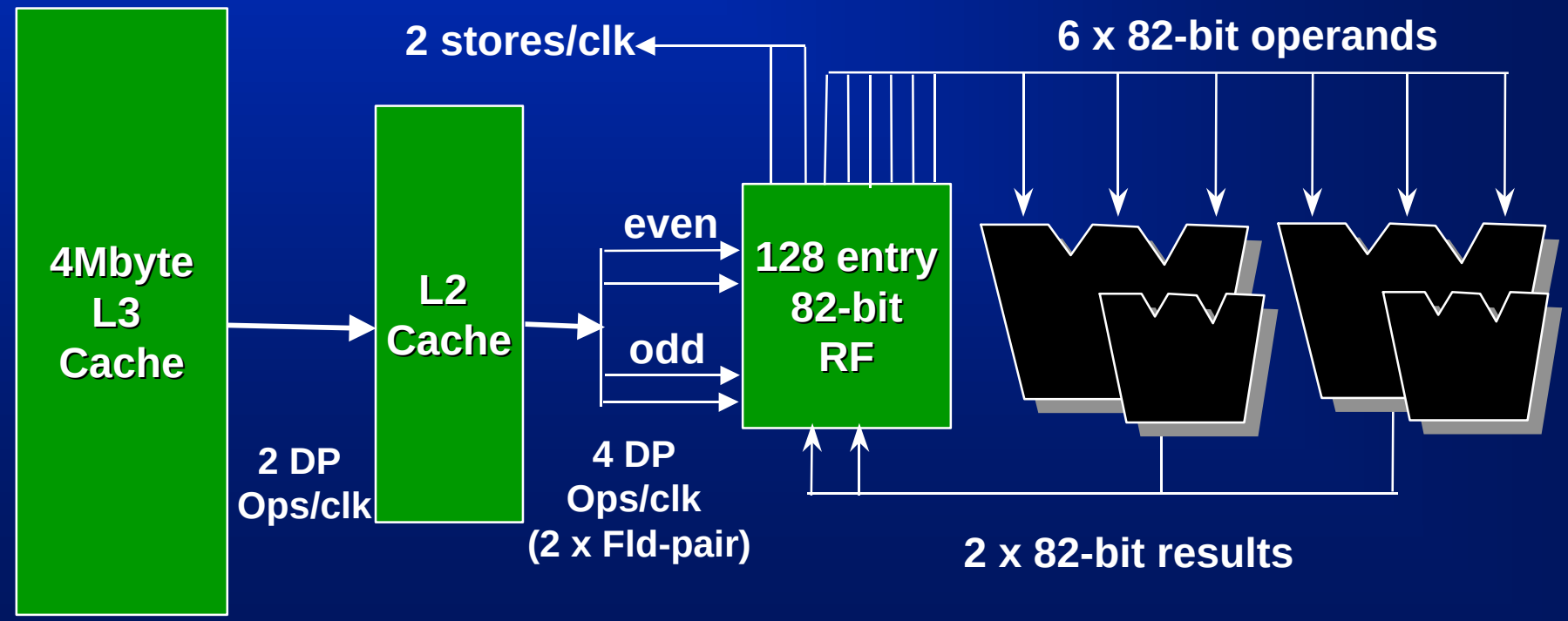
- Control Speculation support requires minimal hardware
  - Computed memory exception delivered with data as tokens (NaTs)
  - NaTs propagate through subsequent executions like source data
- Data Speculation enabled efficiently via ALAT structure
  - 32 outstanding advanced loads
  - Indexed by reg-ids, keeps partial physical address tag
- 0 clk checks: dependent use can be issued in parallel with check



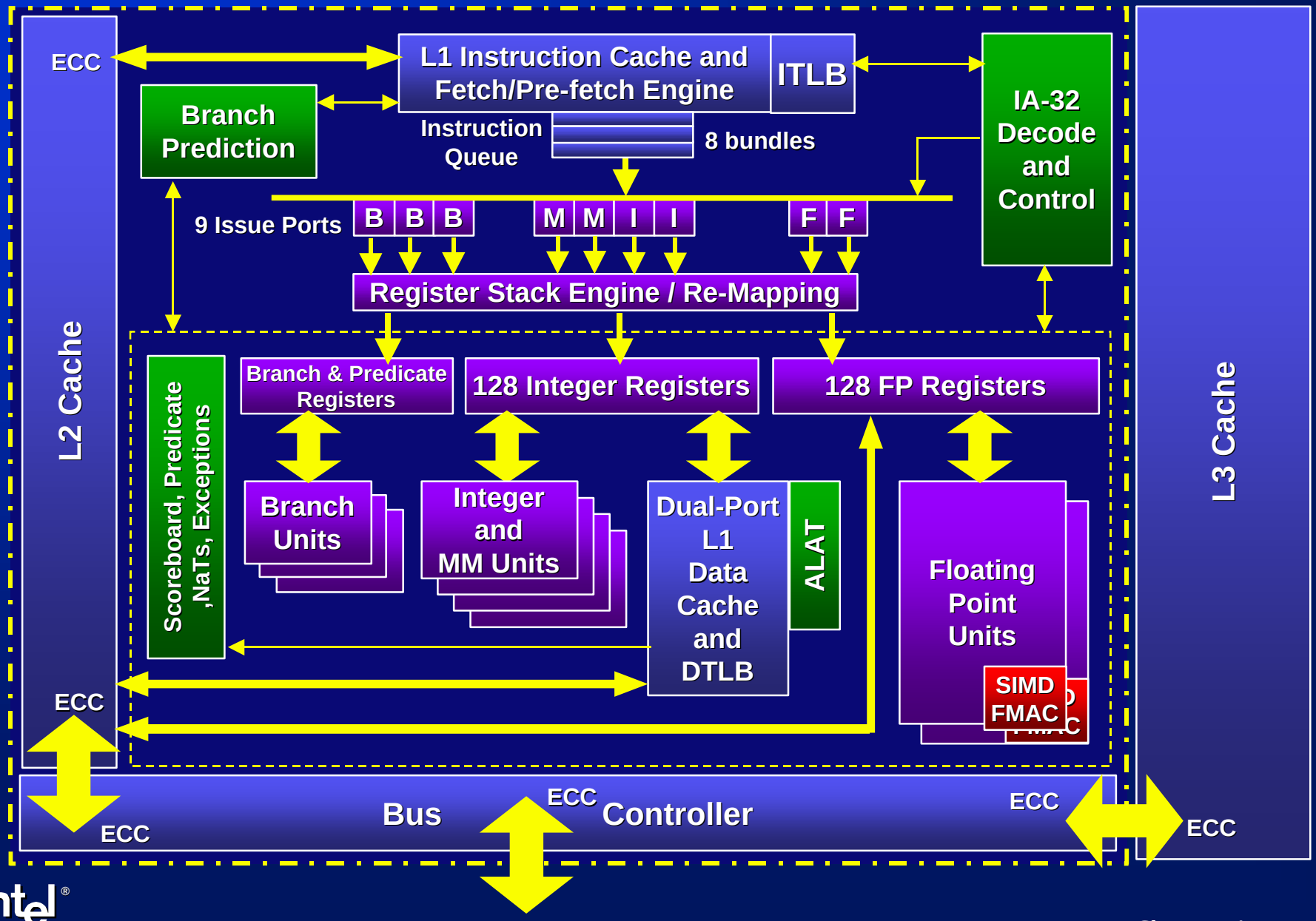
Efficient elimination of memory bottlenecks

Floating Point Features

- Native 82-bit hardware provides support for multiple numeric models
- 2 Extended precision pipelined FMACs deliver 4 EP / DP FLOPs/cycle
- Balanced with plenty of operand bandwidth from registers / memory
- Tuned for 3D graphics: 2 Additional SP FMACs deliver 8 SP FLOPs/cycle; Software divide allows SW pipelining for high throughput;
- FPU hardware used for twin Integer multiply-add (>1000 RSA decrypts/sec)



Intel® Itanium™ Processor Block Diagram



Itanium™ Processor Core Summary

- State-of-the-Art processor for Servers and Workstations
  - Combines High Performance with 64-bit addressing, Reliability features for Mission critical Applications, & full iA-32 compatibility in hardware
- Highly parallel and deeply pipelined hardware at 800Mhz
  - 6-wide, 10-stage pipeline at 800Mhz on 0.18 µ process
- EPIC technology increases Instruction Level Parallelism (ILP)
  - Speculation, Predication, Explicit Parallelism, Register Stacking, Rotation, & Branch/Memory Hints maximize hardware-software synergy
- Dynamic features enable high-throughput on compiled schedule
  - Register scoreboard, non-blocking caches, Decoupled instruction prefetch & aggressive branch prediction
- Supercomputer-level FP (3.2 GFLOPs) for technical workstations